



IGNITED MINDS
Journals

*International Journal of
Information Technology
and Management*

*Vol. VIII, Issue No. XII,
May-2015, ISSN 2249-4510*

**A RESEARCH UPON EFFECTIVE APPROACHES
AND APPLICATION OF DATA MINING FOR
DECISION MAKING IN HEALTH CARE SYSTEM**

AN
INTERNATIONALLY
INDEXED PEER
REVIEWED &
REFEREED JOURNAL

A Research upon Effective Approaches and Application of Data Mining For Decision Making In Health Care System

Sujeet Varshney

Research Scholar, Maharishi Dayanand University, Rajasthan

Abstract – Data Mining is one of the most motivating areas of research that is become increasingly popular in health organization. Data Mining plays an important role for uncovering new trends in healthcare organization which in turn helpful for all the parties associated with this field. This survey explores the utility of various Data Mining techniques such as classification, clustering, association, regression in health domain. In this paper, we present a brief introduction of these techniques and their advantages and disadvantages. This survey also highlights applications, challenges and future issues of Data Mining in healthcare.

Healthcare presents unique challenges for the architect of a data warehouse. Integrated health systems are shifting its focus away from the acute care setting and moving towards cross-continuum care management. Improving healthcare quality while reducing costs requires the elimination of unnecessary variation in the care process. This paper describes the lessons learned during the business case development for the project. Topics include establishing the need for a data warehouse, understanding data warehousing in healthcare, justifying the cost of a data warehouse, building the team, and setting achievable goals.

In this paper, we have focused to compare a variety of techniques, approaches and different tools and its impact on the healthcare sector. The goal of data mining application is to turn that data are facts, numbers, or text which can be processed by a computer into knowledge or information. The main purpose of data mining application in healthcare systems is to develop an automated tool for identifying and disseminating relevant healthcare information.

Tendency for data mining application in healthcare today is great, because healthcare sector is rich with information, and data mining is becoming a necessity. Healthcare organizations produce and collect large volumes of information on daily basis. Use of information technologies allows automatization of processes for extraction of data that help to get interesting knowledge and regularities, which means the elimination of manual tasks and easier extraction of data directly from electronic records, transferring onto secure electronic system of medical records which will save lives and reduce the cost of the healthcare services, as well and early discovery of contagious diseases with the advanced collection of data. Data mining can enable healthcare organizations to predict trends in the patient conditions and their behaviors, which is accomplished by data analysis from different perspectives and discovering connections and relations from seemingly unrelated information. Raw data from healthcare organizations are voluminous and heterogeneous. They need to be collected and stored in the organized forms, and their integration enables forming of hospital information system. Healthcare data mining provides countless possibilities for hidden pattern investigation from these data sets. These patterns can be used by physicians to determine diagnoses, prognoses and treatments for patients in healthcare organizations.

-----X-----

INTRODUCTION

Data Mining is one of the most vital and motivating area of research with the objective of finding meaningful information from huge data sets. In present era, Data Mining is becoming popular in healthcare

field because there is a need of efficient analytical methodology for detecting unknown and valuable information in health data. In health industry, Data Mining provides several benefits such as detection of the fraud in health insurance, availability of medical solution to the patients at lower cost, detection of

causes of diseases and identification of medical treatment methods. It also helps the healthcare researchers for making efficient healthcare policies, constructing drug recommendation systems, developing health profiles of individuals etc.. The data generated by the health organizations is very vast and complex due to which it is difficult to analyze the data in order to make important decision regarding patient health. This data contains details regarding hospitals, patients, medical claims, treatment cost etc. So, there is a need to generate a powerful tool for analyzing and extracting important information from this complex data. The analysis of health data improves the healthcare by enhancing the performance of patient management tasks. The outcome of Data Mining technologies are to provide benefits to healthcare organization for grouping the patients having similar type of diseases or health issues so that healthcare organization provides them effective treatments. It can also be useful for predicting the length of stay of patients in hospital, for medical diagnosis and making plan for effective information system management. Recent technologies are used in medical field to enhance the medical services in cost effective manner. Data Mining techniques are also used to analyze the various factors that are responsible for diseases for example type of food, different working environment, education level, living conditions, availability of pure water, health care services, cultural, environmental and agricultural factors as shown in Figure 1.

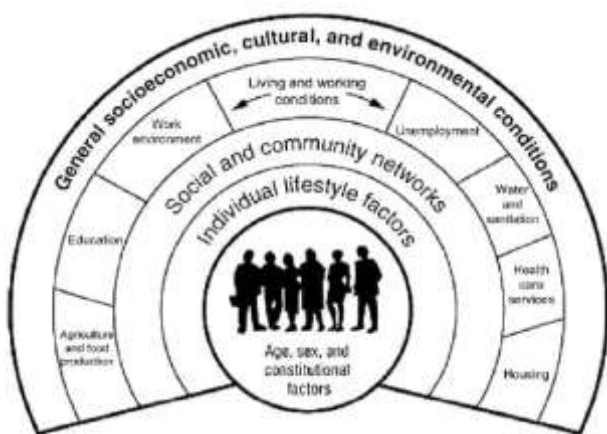


Figure 1. Factors Responsible for Diseases.

Healthcare organizations today are capable of generating and collecting a large amounts of data. This increase in volume of data requires automatic way for these data to be extracted when needed. With the use of data mining techniques it is possible to extract interesting and useful knowledge and regularities.

Knowledge acquired in this manner, can be used in appropriate area to improve work efficiency and enhance quality of decision making process. Above stated points that there is a great need for new generation of computer theories and tools to help people with extracting useful information from constantly growing volume of digital data. Information

technologies are being increasingly implemented in healthcare organizations in order to respond to the needs of doctors in their daily decision making activities. Data mining tools can be very useful to control limitations of people such as subjectivity or error due to fatigue, and to provide indications for the decision-making processes. The essence of data mining is in the identification of relations, patterns and models that provide support for predictions and of decision making process for diagnoses and treatment planning. These models can be called predictive, and they are being integrated into information systems of hospitals as a models for decision making, reducing the subjectivity and decision making time. In addition, the use of information technology in healthcare enables comprehensive management of medical knowledge and its secure exchange between recipients and providers of healthcare services. Widespread use of information technology enables the elimination of manual tasks of data extraction from charts or filling of specialized questionnaires, extraction of data directly from electronic records, transfer on secure electronic system of medical records that will save lives and reduce the cost of health care, early detection of infectious diseases with advanced collection of data etc. Retrieval of information with the help of computers can help the quality of decision making and avoiding human errors.

When there is a large volume of data that needs to be classified, decision making by people is usually poor. Data mining represents the process of analyzing raw data with the help of computer and extraction of their meaning. It is frequently defined as discovering previously unknown and potentially useful information from large volume of (unstructured) data. Thanks to this technique, it is possible to predict trends and customer behavior and thus provide the organization's business success. This is accomplished by data analysis from various perspectives and finding the connections and relations between mutually unconnected information. In the process of data mining previously unknown trends and patterns from a database of historical information are being discovered and that information is being converted into significant business solutions.

Data mining can be defined as the process of finding previously unknown patterns and trends in databases and using that information to build predictive models. Alternatively, it can be defined as the process of data selection and exploration and building models using massive data stores to uncover previously unknown patterns. Data mining is an analytic process designed to explore large amounts of data in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. Data mining is an automated approach for discovering or inferring hidden patterns or knowledge buried in data. 'Hidden' means patterns that are not made

apparent through casual observation. Data Mining is an interdisciplinary field that combines artificial intelligence, computer science, machine learning, database management, data visualization, mathematic algorithms, and statistics. Data Mining is a technology for knowledge discovery in databases (KDD). This technology provides different methodologies for decision making, problem solving, analysis, planning, diagnosis, detection, integration, prevention, learning and innovation. Data mining is a variety of techniques such as neural networks, decision trees or standard statistical techniques to identify nuggets of information or decision making knowledge in bodies of data, and extracting these in such a way that they can be put to use in areas such as decision support, prediction, forecasting, and estimation.

Electronic medical records are becoming more ubiquitous in day-to-day clinical practice. They capture clinical data, store in personal database as well as mirror it in local and regional databases. Data capture, storage, retrieval and display are all performed. They also allow display of alerts, warnings; guide a clinician through a clinical protocol by way of workflow, and online transactional processing where "intelligent" data display is made through running structured queries using SQL, etc.

Unfortunately, all this data residing in a RDBMS is good enough for basically the following purposes with respect to improving patient care.

1. Display against a period of time allowing for better visualization of the patient's clinical condition
2. Potentially life-saving alerts/warnings about a patient based on the clinical information collected about the patient They are not able to support either evidence based medicine or outcomes analysis directly. To perform these tasks, it is necessary to have a data warehouse or at least an appropriate custom-built interface for the same. Although running specially designed queries may be able to accomplish this task, the payoff is that the user needs to correctly design them and retrieving results proves to a slow process as the data is usually not "analysis-ready".

It does however hold the potential to unleash a revolutionizing wealth of information regarding disease processes, disease progression, best method of treatment, optimizing costs while maximizing efficiency, etc. Currently, the way to make this possible is to use online analytical processing by way of using data warehousing and data mining. Data mining, functionally, is the process of discovering interesting knowledge from large amounts of data stored in various data repositories like databases or data warehouses. The process involves integration of

techniques like database technology, statistics, artificial intelligence, high performance computing, data visualization, image/signal processing, and spatial data analysis. By performing this process, interesting knowledge, patterns and high-level information can be extracted, viewed and browsed from multiple angles. The knowledge so discovered can be applied to decision-making, process control, information management, and query processing.

Managing data in healthcare organizations has become a challenge as a result of healthcare managers having considerable differences in objectives, concerns, priorities and constraints. The planning, management and delivery of healthcare services included the manipulation of large amounts of health data and the corresponding technologies have become increasingly embedded in all aspects of healthcare. Information is one of the most factors to an organization success that executive managers or physicians would need to base their decisions on, during decision making.. Healthcare organizations typically deal with large volumes of data containing valuable information about patients, procedures, treatments and etc. These data are stored in operational databases that are not useful for decision makers or executives.

A majority of database management systems used in these organizations execute online transaction processing (OLTP) direct answer to queries at the executive level, such as the what-if and what-next type queries. Decision makers at executive level would like to quickly analyze existing health data and in time to aid in the decision making process. However, stand-alone databases cannot provide such information quickly and efficiently. The concept of data warehousing provides a powerful solution for data integration and information access problems.

DATA MINING

Data Mining came into existence in the middle of 1990's and appeared as a powerful tool that is suitable for fetching previously unknown pattern and useful information from huge dataset. Various studies highlighted that Data Mining techniques help the data holder to analyze and discover unsuspected relationship among their data which in turn helpful for making decision. In general, Data Mining and Knowledge Discovery in Databases (KDD) are related terms and are used interchangeably but many researchers assume that both terms are different as Data Mining is one of the most important stages of the KDD process. According to Fayyad et al., the knowledge discovery process are structured in various stages whereas the first stage is data selection where data is collected from various sources, the second stage is pre - processing of the selected data , the third stage is the transformation of

the data into appropriate format for further processing, the fourth stage is Data Mining where suitable Data Mining technique is applied on the data for extracting valuable information and evaluation is the last stage as shown in Figure 2.

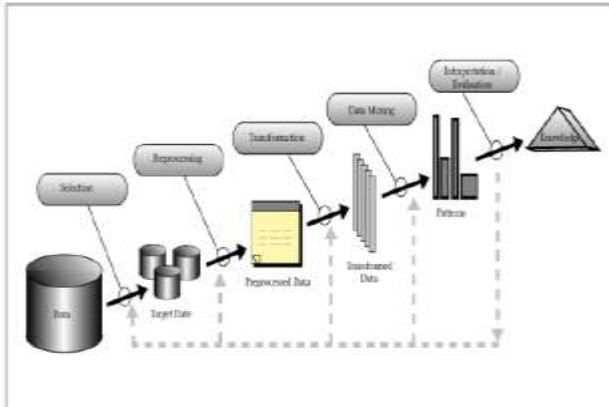


Figure 2. Stages of Knowledge Discovery Process.

Skills and knowledge are essential requirement for performing the Data Mining task because the success and failure of Data Mining projects is greatly dependent on the person who are managing the process due to unavailability of standard framework. The CRISP-DM (CRoss Industry Standard Process for Data Mining) provides a framework for carrying out Data Mining activities. CRISP-DM divides the data mining task into 6 phases. The first phase is the understanding of the business activities while the data for carrying out business activities are collected and analyzed in the second phase. Data pre-processing and modelling is done in the third and fourth phase respectively. Fifth phase evaluates the model and last phase is responsible for deployment of the construed model.

Nowadays there is huge amount of data stored in real-world databases and this amount continues to grow fast as it creates both an opportunity and a need for semi-automatic methods that discover the hidden knowledge in such database. If such knowledge discovery activity is successful, discovered knowledge can be used to improve the decision making process of an organization. For instance data about a hospital's patient might contain interesting knowledge about which kind of patient is more likely to develop a given disease. This knowledge can lead to better diagnosis and treatment for future patients. Data mining and knowledge discovery is the name often used to refer to a very interdisciplinary field, which consists of using methods of several research areas to extract knowledge from real world data sets. There is a distinction between the terms data mining and knowledge discovery; the term data mining refers to the core steps of a broader process called knowledge discovery in database. In addition to the data mining step which actually extracts knowledge from data, the knowledge discovery process includes several preprocessing and post processing steps. The goal of data preparation methods is to transform the data to

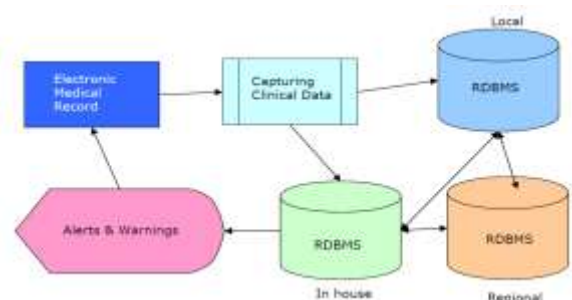
facilitate the application of a given data mining algorithms, where the goal of knowledge refinement methods is to validate and refine discovered knowledge. The knowledge discovery is both iterative and interactive. It is iterative because the output of each step is often feedback to previous steps and typically many iterations of this process are necessary to extract high quality knowledge from data. It is interactive because the user or more precisely an expert in the application domain should be involved in this loop to help in data preparation, discovered-knowledge validation and refinement.



Fig 3: Data Mining Architecture.

CURRENT STATUS

Electronic medical records capable of capturing clinical data, storing them locally (i.e. the clinician's own database), in-house (i.e., in the same organization like clinic, department or hospital), and regionally (i.e. in the same geographical area), capable of displaying data on request, alerts that are rule-based and patient-specific (display an alert if this patient's systolic blood pressure comes down below 60 mmHg or fasting blood sugar is more than 110 mg/dL on three consecutive days, etc.), warnings that have been pre-set (display a warning if any patient allergic to penicillin group of drugs and is suffering from rheumatic fever with ASO titer more than 200 Todd units, or a contra-indicated drug is prescribed, or two interacting drugs are prescribed concomitantly, etc.), following clinical protocols and performing other OLTP functions. The relevant software architecture is as follows.



DATAMINING APPROCHES IN HEALTH CARE

Data mining applications can be developed to evaluate the effectiveness of medical treatments. By comparing and contrasting causes, symptoms, and courses of treatments, data mining can deliver an analysis of which courses of action prove effective, for example the outcomes of patient groups treated with different drug regimens for the same disease or condition can be compared to determine which treatments work best and are most cost-effective. Other data mining applications related to treatments include associating the various side-effects of treatment, collating common symptoms to aid diagnosis, determining the most effective drug compounds for treating sub-populations that respond differently from the mainstream population to certain drugs and determining proactive steps that can reduce the risk of affliction. future needs of individuals to improve their level of satisfaction. These applications also can be used to predict other products that a healthcare customer is likely to purchase, whether a patient is likely to comply with prescribed treatment or whether preventive care is likely to produce a significant reduction in future utilization.

There are various data mining techniques available with their suitability dependent on the domain application. Statistics provide a strong fundamental background for quantification and evaluation of results. However, algorithms based on statistics need to be modified and scaled before they are applied to data mining. We now describe a few Classification data mining techniques with illustrations of their applications to healthcare.

A. Rule set classifiers:

Complex decision trees can be difficult to understand, for instance because information about one class is usually distributed throughout the tree. C4.5 introduced an alternative formalism consisting of a list of rules of the form "if A and B and C and ... then class X", where rules for each class are grouped together. A case is classified by finding the first rule whose conditions are satisfied by the case; if no rule is satisfied, the case is assigned to a default class IF conditions.

THEN conclusion - This kind of rule consists of two parts. The rule antecedent (the IF part) contains one or more conditions about value of predictor attributes whereas the rule consequent (THEN part) contains a prediction about the value of a goal attribute. An accurate prediction of the value of a goal attribute will improve decision-making process. IF-THEN prediction rules are very popular in data mining; they represent discovered knowledge at a high level of abstraction.

B. Decision Tree algorithms:

Decision tree include CART (Classification and Regression Tree), ID3 (Iterative Dichotomized 3) and C4.5. These algorithms differ in selection of splits, when to stop a node from splitting, and assignment of class to a non-split node. CART uses Gini index to measure the impurity of a partition or set of training tuples. It can handle high dimensional categorical data. Decision Trees can also handle continuous data (as in regression) but they must be converted to categorical data. We will refer to a row as a data instance. The data set contains three predictor attributes, namely Age, Gender, Intensity of symptoms and one goal attribute, namely disease whose values (to be predicted from symptoms) indicates whether the corresponding patient have a certain disease or not.

C. Neural Network Architecture :

The architecture of the neural network used in this study is the multilayered feed-forward network architecture with 20 input nodes, 10 hidden nodes, and 10 output nodes. The number of input nodes is determined by the finalized data; the number of hidden nodes is determined through trial and error and the number of output nodes is represented as a range showing the disease classification. The most widely used neural-network learning method is the BP algorithm.

Learning in a neural network involves modifying the weights and biases of the network in order to minimize a cost function. The cost function always includes an error term a measure of how close the network's predictions are to the class labels for the examples in the training set.

Additionally, it may include a complexity term that reacts to a prior distribution over the values that the parameters can take. Neural networks have been proposed as useful tools in decision making in a variety of medical applications. Neural networks will never replace human experts but they can help in screening and can be used by experts to double-check their diagnosis. In general, results of disease classification or prediction task are true only with a certain probability.

D. Neuro-Fuzzy:

Stochastic back propagation algorithm is used for the construction of fuzzy based neural network. The steps involved in the algorithm are as follows: First, initialize weights of the connections with random values. Second for each unit compute net input value, output value and error rate. Third, to handle uncertainty for each node, certainty measure (c) for each node is calculated. Based on the certainty measure the decision is made.

DATA MINING APPLICATIONS IN HEALTHCARE SECTOR

Healthcare industry today generates large amounts of complex data about patients, hospital resources, disease diagnosis, electronic patient records, medical devices etc. Larger amounts of data are a key resource to be processed and analyzed for knowledge extraction that enables support for cost-savings and decision making. Data mining applications in healthcare can be grouped as the evaluation into broad categories,

Treatment effectiveness :

Data mining applications can develop to evaluate the effectiveness of medical treatments. Data mining can deliver an analysis of which course of action proves effective by comparing and contrasting causes, symptoms, and courses of treatments.

Healthcare management :

Data mining applications can be developed to better identify and track chronic disease states and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims to aid healthcare management. Data mining used to analyze massive volumes of data and statistics to search for patterns that might indicate an attack by bio-terrorists.

Effective management of Hospital resource:

Data mining provides support for constructing a model for managing the hospital resources which is an important task in healthcare. Using data mining, it is possible to detect the chronic disease and based on the complication of the patient disease prioritize the patients so that they will get effective treatment in timely and accurate manner. Fitness report and demographic details of patients is also useful for utilizing the available hospital resources effectively. An automated tool using data mining is proposed by J.Alapont et al., for managing hospital resources such as physical and human resources. Group Health Cooperative provides various healthcare services at lower cost using data mining techniques. It is a non-profit organization of healthcare that offers patients to online access their medical information, online fill the prescription form and allow safe exchanging of e-mail with the healthcare provider. Seton Medical centre also used data mining to enhance the healthcare quality, provide various details regarding patient's health and reduce admitted duration of the patients in the hospitals. With the help of data mining Blue Cross provide a system for managing the diseases efficiently and improve the results and lower the cost of expenditure. Sierra Health Centre provides guidelines for treatment, managing the cost of treatment and detects the areas for improving the health quality using data mining.

Hospital Ranking:

Different data mining approaches are used to analyze the various hospital details in order to determine their ranks. Ranking of the hospitals are done on the basis of their capability to handle the high risk patients. The hospital with higher rank handles the high risk patient on its top priority while the hospital with lower rank does not consider the risk factor.

Better Customer Relation:

Data Mining helps the healthcare institute to understand the needs, preferences, behavior, patterns and quality of their customer in order to make better relation with them. Using Data Mining, Customer Potential Management Corp. develops an index represent the utilization of Consumer healthcare. This index helps to detect the influence of customer towards particular healthcare service.

Hospital Infection Control:

A system for inspection is constructed using data mining techniques to discover unknown or irregular patterns in the infection control data. Association rules are used to produce unexpected and interesting information from the public surveillance and hospital control data. To control the infection in the hospitals, this information is reviewed further by an Expert.

Smarter Treatment Techniques:

Using Data Mining, physicians and patients can easily compare among different treatments technique. They can analyze the effectiveness of available treatments and find out which technique is better and cost effective. Data Mining also helps them to identify the side effects of particular treatment, to make appropriate decision to reduce the hazard and to develop smart methodologies for treatment.

Improved Patient care:

Large amount of data is collected with the advancement in electronic health record. Patient data which is available in digitized form improve the healthcare system quality. In order to analyze this massive data, a predictive model is constructed using data mining that discover interesting information from this huge data and make decision regarding the improvement of healthcare quality. Data mining helps the healthcare providers to identify the present and future requirements of patients and their preferences to enhance their satisfaction levels. Milley has also recommended that data mining are useful to determine the requirement of particular patients for enhancing the services provided by healthcare organization. Hallick has suggested that Data mining techniques are helpful to provide the information to patient regarding various diseases and their prevention. Kolar has

identified that healthcare organization used data mining techniques for patient grouping.

Decrease Insurance Fraud:

Healthcare insurer develops a model to detect the fraud and abuse in the medical claims using data mining techniques. This model is helpful for identifying the improper prescriptions, irregular or fake patterns in medical claims made by physicians, patients, hospitals etc. US taxpayers also reported to lost hundred dollars in 1997 due to fraudulent in the hospitals bill. ReliaStar financial corp. has improved the annual savings by 20% by detected the fraud and abuse. Doctor's prescriptions and treatment materials are produced large amount of data. Utah Bureau of Medicaid Fraud used this data to discover hidden and useful information in order to detect fraud. Australian Health Insurance Commission has also mined the huge data and reported millions of dollars of annual saving. Texas Medicaid Fraud and Abuse Detection System have also used data mining techniques to discover the fraud and abuse and saved million dollars in 1998.

Recognize High-Risk Patients:

American Healthways system construct a predictive model using data mining to recognize the patients having high risk. The main concern of this system is to handle the diabetic patients, improve their health quality and also offers cost savings services to the patient. Using Predictive model, healthcare provider recognize the patient which require more concern as compare to other patients.

Health Policy Planning:

Data mining play an important role for making effective policy of healthcare in order to improve the health quality as well as reducing the cost for health services. COREPLUS and SAFS models were developed using data mining techniques to analyze the results of medical care services provided by hospitals and treatment cost.

CONCLUSION

This paper aimed to compare the different data mining application in the healthcare sector for extracting useful information. The prediction of diseases using Data Mining applications is a challenging task but it drastically reduces the human effort and increases the diagnostic accuracy. Developing efficient data mining tools for an application could reduce the cost and time constraint in terms of human resources and expertise. Exploring knowledge from the medical data is such a risk task as the data found are noisy, irrelevant and massive too.

Data mining helps professionals discover these patterns and put them to work. As models are based directly on history, they represent the ultimate in evidence-based care. But technology is no panacea, and professional, ethical and practical issues must be addressed. Decisions must rest with the healthcare professionals, not the information systems experts.

In a clinical setting, the first step is to capture, authenticate, validate, store and retrieve the data in the proper manner. This is already being done through electronic medical records. The next is to unearth the knowledge that lies "hidden" within the captured data.

The main goal of this survey was to identify the most common data mining algorithms, implemented in modern Healthcare Decision Support Systems, and evaluate their performance on several medical datasets.

We can also conclude that there is no single data mining techniques which give consistent results for all types of healthcare data. The performance of data mining techniques depends on the type of dataset that we have taken for doing experiment. So, we can use hybrid or integrated Data Mining technique such as fusion of different classifiers, fusion of clustering with classification or association with clustering or classification etc. for achieving better performance.

For effective utilization of data mining in health organizations there is a need of enhance and secure health data sharing among different parties. Some propriety limitations such as contractual relationships among researcher and health care organization are mandatory to overcome the security issues.

REFERENCES

- A.A. Freitas, "Understanding the crucial role of attribute interaction in data mining," *Artificial Intelligence Review*, vol. 16, pp. 177-199, 2001.
- Arvind Sharma and P.C. Gupta —Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool|| *International Journal of Communication and Computer Technologies* Volume 01 – No.6, Issue: 02 September 2012.
- Bushinak, H., Abdel Gaber, S., & Al Sharif, F. K. (2011). Recognizing The Electronic Medical Record Data From Unstructured Medical Data Using Visual Text Mining Techniques Prof. Hussain Bushinak. (*IJCSIS*) *International Journal of Computer Science and Information Security*, Vol. 9, No. 6 , 25-35.

- C.McGregor, C. Christina and J. Andrew, "A process mining driven framework for clinical guideline improvement in critical care", Learning from Medical Data Streams 13th Conference on Artificial Intelligence in Medicine (LEMEDS). <http://ceur-ws.org>, vol. 765, (2012).
- H.C. Koh and G. Tan, "Data Mining Application in Healthcare", Journal of Healthcare Information Management, vol. 19, no. 2, (2005).
- J.Dych (2000) "E-Data Turning Data into Information with Data Warehousing", Addison-Wesley, Reading.
- K.Srinivas , B. Kavitha Rani and Dr. A. Govrdhan, —Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks|| International Journal on Computer Science and Engineering (2010).
- Koh, H. C., & Tan, G. (2005). Data Mining Applications in Healthcare. Journal of Healthcare Information Management -Vol. 19, No. 2 , 64-72.
- Kolar, H.R. (2001). Caring for healthcare. Health Management Technology, 22(4), 46-47.
- Laura Hadley, (2002) "Developing a Data Warehouse Architecture".
- N. Padhy, "THE SURVEY OF DATA MINING APPLICATIONS AND FEATURE SCOPE Pragnyaban Mishra, Neelamadhab Padhy," 2012.
- R.D. Canlas Jr., "Data Mining in Healthcare: Current Applications and Issues", (2009).
- Silver, M. Sakata, T. Su, H.C. Herman, C. Dolins, S.B. & O'Shea, M.J. (2001). Case study: how to apply data mining techniques in a healthcare data warehouse. Journal of Healthcare Information Management, 15(2), 155-164.
- T.-H. Chen and C.-W. Chen, "Application of data mining to the spatial heterogeneity of foreclosed mortgages," Expert Systems with Applications, vol. 37, pp. 993-997, 2010.